

УДК 004.93

Ст. викладач Темнікова О.Л., к.т.н., доцент Темніков В.О.,
магістрант Конфорович І.В.

Національний технічний університету України
«Київський політехнічний інститут»

АЛГОРИТМ ПОБУДОВИ СИСТЕМИ ПАРАМЕТРИЗАЦІЇ ДЛЯ АВТЕНТИФІКАЦІЇ ОПЕРАТОРІВ ЗА ГОЛОСОМ

Abstract

*Olena L.Temnikova, senior lecturer, Volodymyr O.Temnikov,
assoc.prof., Ph.D., Igor V.Konforovich, student
Algorithm for development of parameterization system for voice
authentication of operators*

This article proposes algorithm for development of parameterization system for voice authentication of ergatic systems' subjects (operators) and for calculation of keprstral coefficients as informative parameters of voice signals for further authentication.

Вступ

Одним із шляхів підвищення безпеки систем управління ергатичних систем (ЕС) є неперервний контроль за діями операторів як суб'єктів ЕС у процесі виконання їх функціональних обов'язків під час робочої зміни, що дозволяє істотно знизити відсоток виникнення аварій та аварійних ситуацій через людський фактор.

Через це актуальною є розробка системи контролю доступу (СКД) операторів до систем управління ЕС, що забезпечує вирішення таких задач як автентифікація операторів (для запобігання несанкціонованого доступу до систем управління) і моніторинг їх психофізіологічного стану [1].

Постановка задачі

Автентифікацію операторів пропонується проводити в автоматичному режимі за голосом, аналізуючи параметри отриманих у процесі аудіообміну між операторами мовних відрізків, що являють собою неперервну мову.

Метою роботи є розробка алгоритму побудови однієї з підсистем СКД – системи параметризації мовних сигналів операторів. Система

параметризації розробляється для випадку, коли аудіообмін між операторами відбувається з використанням встановленої фразеології.

Розроблену систему буде зокрема застосовано для контролю роботи авіадиспетчерів, які використовують певну фразеологію в процесі аудіообміну з повітряними і наземними авіаційними станціями [2].

Алгоритм побудови системи параметризації

Розроблений алгоритм побудови системи параметризації включає такі етапи: сегментацію, пошук слів в неперервній мові (виділення слів із завчасно складеного словника) і параметризацію лексичних елементів для подальшого розпізнавання оператора (за виділеним словом проводимо автентифікацію оператора за дикторозалежними параметрами).

Обробка мовних сигналів на вказаних етапах проводиться зі застосуванням таких методів досліджень: сегментація – з використанням вейвлет-перетворень, пошук слів у неперервному тексті – на основі прихованих марковських моделей, розпізнавання (автентифікація) оператора – на основі структурних елементів класичної теорії розпізнавання образів [3].

Розрахунок кепстральних коефіцієнтів мовного сигналу

При розпізнаванні оператора за голосом в якості образу, що аналізується, виступає часова залежність, що описує мову конкретної особи. Складність задачі розпізнавання особи за голосом полягає в знаходженні інформативних параметрів, що гарантовано характеризують її, і у розробці методів ефективного виділення цих параметрів для обробки в оптимальному математичному представленні. „Вдале” виділення інформативних параметрів є основою для прийняття системою автентифікації достовірного рішення.

В якості інформативних параметрів мовних сигналів для автентифікації операторів авторами пропонується використовувати кепстральні коефіцієнти, які визначаються в такій послідовності:

перший крок: застосування віконної функції Ханна [4]

$$\omega(n) = 0.5 \left(1 - \cos \left(2\pi \frac{n}{N} \right) \right), \quad 0 \leq n \leq N,$$

де n – номер відліку, N – загальна кількість відліків;

другий крок: розрахунок коефіцієнтів лінійного передбачення автокореляційним методом [5];

третій крок: визначення кепстральних коефіцієнтів з коефіцієнтів лінійного передбачення за формулою

$$c(n) = \begin{cases} 0, n < 0 \\ \log_{\varepsilon}(A), n = 0 \\ a_n + \sum_{k=1}^{n-1} \binom{k}{n} c(k) a_{n-k}, 0 < n < p \\ \sum_{k=n-p}^{n-1} \binom{k}{n} c(k) a_{n-k}, n > p \end{cases}$$

де a_i – коефіцієнти лінійного передбачення, $c(i)$ – кепстральні коефіцієнти, p – порядок лінійного передбачення, n – номер кепстрального коефіцієнта, A – похибка лінійного передбачення.

На рис. 1 зображено вихідний мовний сигнал (абсциса – відліки в секундах, ординати - дБ), а також отримані значення (ординати) коефіцієнтів лінійного передбачення і кепстральних коефіцієнтів для фрази «говорите медленнее» по їх номерах (абсциси), що вимовляється диктором, який знаходиться у спокійному психофізіологічному стані.

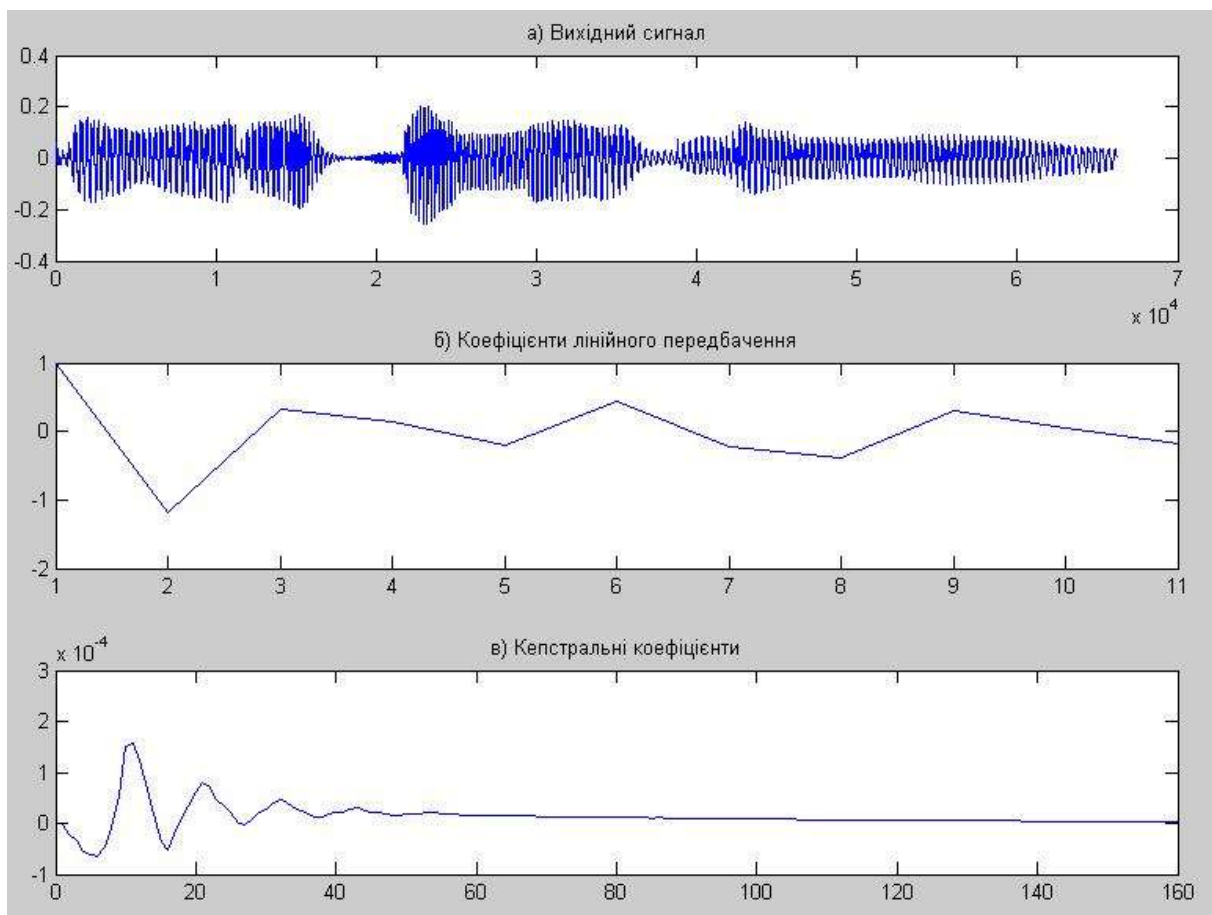


Рис.1. Вихідний мовний сигнал (а), коефіцієнти лінійного передбачення (б) та кепстральні коефіцієнти (в) для фрази «говорите медленнее»

Експерименти, в процесі яких декілька дикторів багато разів вимовляли різні фрази, показали добру повторюваність значень кепстральних коефіцієнтів. В результаті було виявлено, що інформативними для можливості проведення автентифікації є перші 12-20 кепстральних коефіцієнтів.

Кількісними показниками можуть служити коефіцієнти кореляції, розраховані при вимовленні однієї і тієї ж фрази одним і тим же диктором. Наприклад, коефіцієнти кореляції, розраховані для кепстральних коефіцієнтів при вимовленні фрази «говорите медленнее» одним і тим самим диктором 100 разів рівним спокійним голосом, склали: для 160 кепстральних коефіцієнтів – від 0,949 до 0,980, а для 20 кепстральних коефіцієнтів – від 0,99 до 0,999.

Висновки

В статті наведено алгоритм побудови системи параметризації СКД, а також описано результати розрахунку та аналізу кепстральних коефіцієнтів у якості інформативних параметрів для проведення автентифікації операторів як суб'єктів ЕС за голосом в автоматизованій системі контролю доступу.

Література

1. *Темников В.А., Пономаренко Л.В.* Система распознавания личности как основа повышения эффективности систем контроля и управления доступом // Вестник Восточноукраинского национального университета им. В.Даля. - №9(103). – Ч.1. – 2006. – С.64-69
2. *Темников В.А., Темникова Е.Л.* Параметризация автоматического контроля доступа операторов к ресурсам информационных систем по голосу // Вестник Восточноукраинского национального университета им. В.Даля. - №9 (151). – Ч.1. – 2010. – С.143-148
3. *Рамшивили Г. С.* Автоматическое опознавание говорящего по голосу. – М.: Радио и связь, 1981. – 224 с.
4. *Weisstein, Eric W.* Hanning Function // From MathWorld - A Wolfram Web Resource. - <http://mathworld.wolfram.com/HanningFunction.html>
5. *Methods for Speech Recognition - A Project for the Spectral Analysis course ELEC 532.* Rice University, May 1998. - <http://www.clear.rice.edu/elec532/PROJECTS98/speech/cepstrum/cepstrum.html>